In [1]: ►
```python
# Load the Pandas libraries with alias 'pd'
import pandas as pd
# Read data from file 'filename.csv'
# (in the same directory that your python process is based)
# Control delimiters, rows, column names with read_csv (see later)
data = pd.read_csv("C:\\Users\santh\OneDrive\Documents\Spring 2020 Assignme
# Preview the first 5 lines of the loaded data
data.head()
```

Out[1]:

	Car_Name	Year	Selling_Price	Present_Price	Kms_Driven	Fuel_Type	Seller_Type	Transm
0	ritz	2014	3.35	5.59	27000	Petrol	Dealer	l
1	sx4	2013	4.75	9.54	43000	Diesel	Dealer	l
2	ciaz	2017	7.25	9.85	6900	Petrol	Dealer	l
3	wagon r	2011	2.85	4.15	5200	Petrol	Dealer	l
4	swift	2014	4.60	6.87	42450	Diesel	Dealer	l

In [2]: ►
```python
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 301 entries, 0 to 300
Data columns (total 9 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Car_Name       301 non-null    object
 1   Year           301 non-null    int64
 2   Selling_Price  301 non-null    float64
 3   Present_Price  301 non-null    float64
 4   Kms_Driven     301 non-null    int64
 5   Fuel_Type      301 non-null    object
 6   Seller_Type    301 non-null    object
 7   Transmission   301 non-null    object
 8   Owner          301 non-null    int64
dtypes: float64(2), int64(3), object(4)
memory usage: 16.5+ KB
```

In [3]: ▶| `pd.isnull(data).sum()`

Out[3]:
```
Car_Name         0
Year             0
Selling_Price    0
Present_Price    0
Kms_Driven       0
Fuel_Type        0
Seller_Type      0
Transmission     0
Owner            0
dtype: int64
```

In [4]: ▶|
```
data['Kms_Driven'] = data['Kms_Driven'].div(100)
data
```

Out[4]:

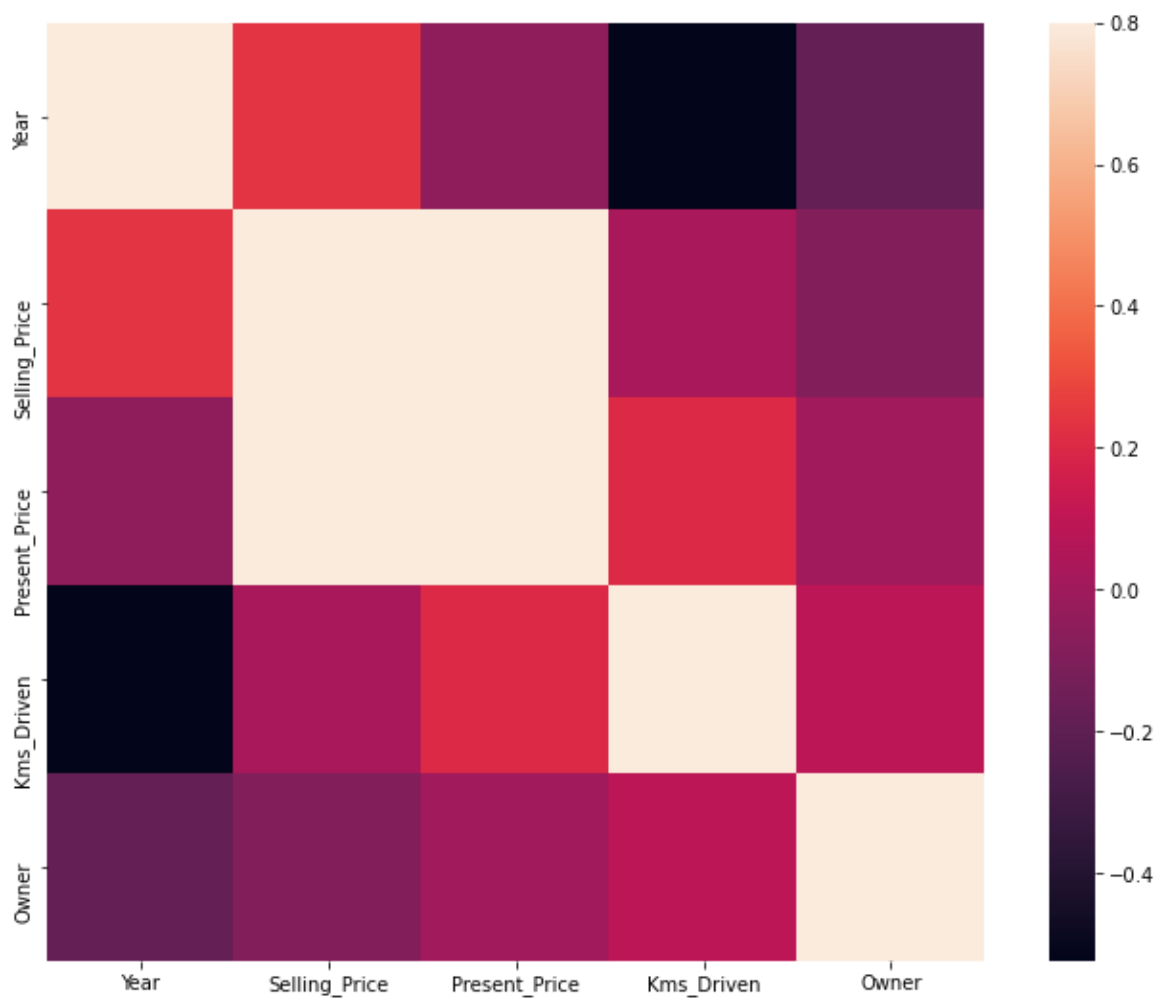|	Car_Name	Year	Selling_Price	Present_Price	Kms_Driven	Fuel_Type	Seller_Type	Trans
0	ritz	2014	3.35	5.59	270.00	Petrol	Dealer	
1	sx4	2013	4.75	9.54	430.00	Diesel	Dealer	
2	ciaz	2017	7.25	9.85	69.00	Petrol	Dealer	
3	wagon r	2011	2.85	4.15	52.00	Petrol	Dealer	
4	swift	2014	4.60	6.87	424.50	Diesel	Dealer	
...	...	...	...	...	...	...	...	
296	city	2016	9.50	11.60	339.88	Diesel	Dealer	
297	brio	2015	4.00	5.90	600.00	Petrol	Dealer	
298	city	2009	3.35	11.00	879.34	Petrol	Dealer	
299	city	2017	11.50	12.50	90.00	Diesel	Dealer	
300	brio	2016	5.30	5.90	54.64	Petrol	Dealer	

301 rows × 9 columns

In [5]: ▶|
```
import numpy as np
numerical_values=data.select_dtypes(include=[np.number])
```

In [6]: ▶
```
corr=numerical_values.corr()
corr
```

Out[6]:

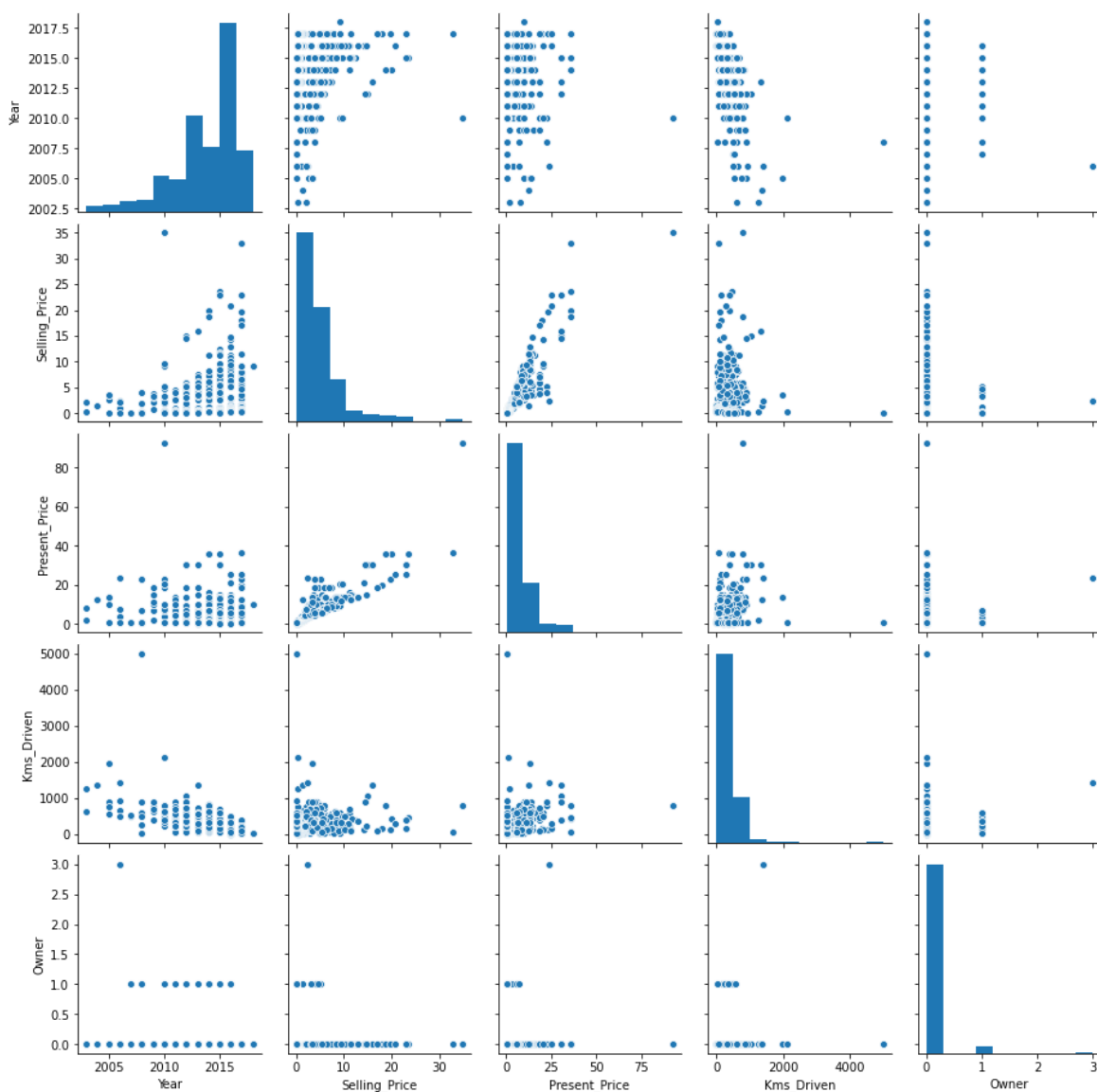|	Year	Selling_Price	Present_Price	Kms_Driven	Owner
Year	1.000000	0.236141	-0.047584	-0.524342	-0.182104
Selling_Price	0.236141	1.000000	0.878983	0.029187	-0.088344
Present_Price	-0.047584	0.878983	1.000000	0.203647	0.008057
Kms_Driven	-0.524342	0.029187	0.203647	1.000000	0.089216
Owner	-0.182104	-0.088344	0.008057	0.089216	1.000000

In [7]:
```python
import seaborn as sns
import matplotlib.pyplot as plt
f,ax=plt.subplots(figsize=(12,9))
display((sns.heatmap(corr,vmax=.8,square=True)).figure)
```
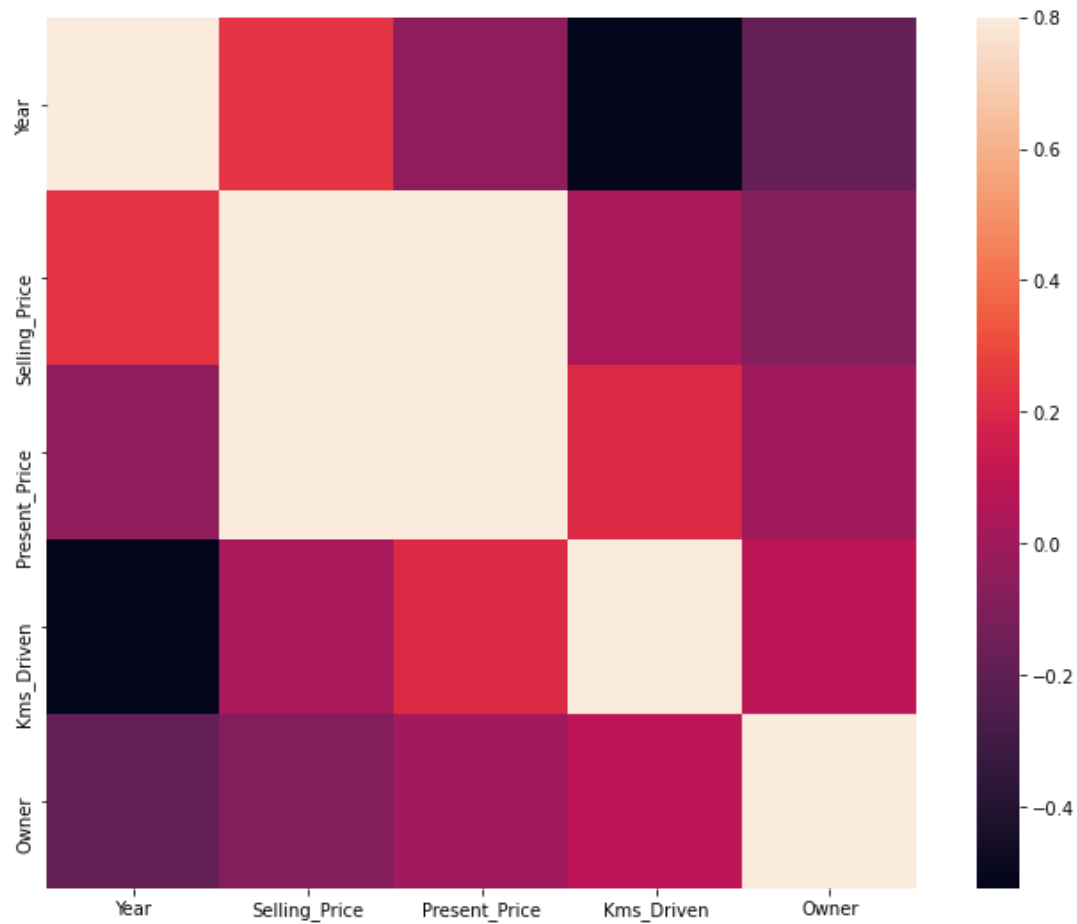
In [8]: ▶|
```python
plt.figure()
sns.pairplot(data)
display(plt.show())
```

<Figure size 432x288 with 0 Axes>

None

In [9]: ▶|
```
data.replace(to_replace=['Manual', 'Automatic', 'Individual', 'Dealer', 'Pe
              value= ['0', '1', '0', '1', '0', '1', '2'],
              inplace=True)
data
```

Out[9]:

	Car_Name	Year	Selling_Price	Present_Price	Kms_Driven	Fuel_Type	Seller_Type	Trans
0	ritz	2014	3.35	5.59	270.00	0	1	
1	sx4	2013	4.75	9.54	430.00	1	1	
2	ciaz	2017	7.25	9.85	69.00	0	1	
3	wagon r	2011	2.85	4.15	52.00	0	1	
4	swift	2014	4.60	6.87	424.50	1	1	
...	...	...	...	...	...	...	...	
296	city	2016	9.50	11.60	339.88	1	1	
297	brio	2015	4.00	5.90	600.00	0	1	
298	city	2009	3.35	11.00	879.34	0	1	
299	city	2017	11.50	12.50	90.00	1	1	
300	brio	2016	5.30	5.90	54.64	0	1	

301 rows × 9 columns

In [10]: ▶|
```python
%matplotlib inline
import matplotlib.pyplot as plt
data.hist(bins=50, figsize=(20,15))
plt.show()
```

c:\users\santh\appdata\local\programs\python\python38-32\lib\site-packages
\pandas\plotting_matplotlib\tools.py:298: MatplotlibDeprecationWarning:
The rowNum attribute was deprecated in Matplotlib 3.2 and will be removed t
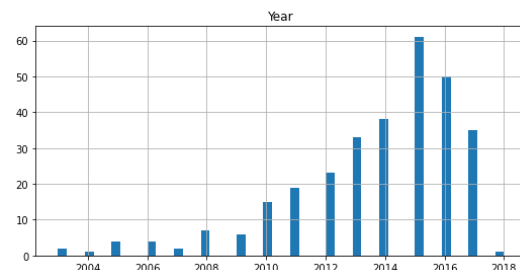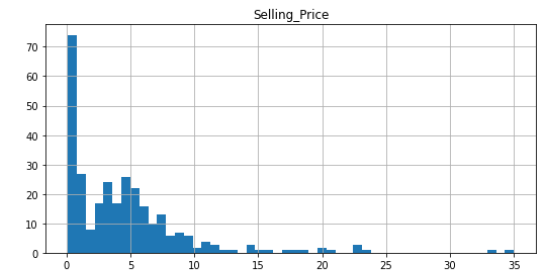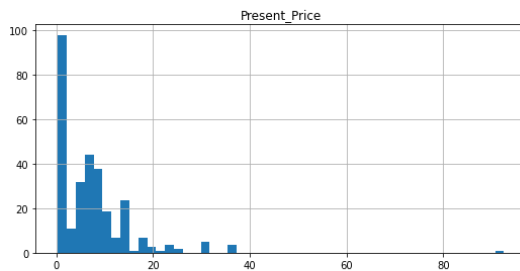wo minor releases later. Use ax.get_subplotspec().rowspan.start instead.
  layout[ax.rowNum, ax.colNum] = ax.get_visible()
c:\users\santh\appdata\local\programs\python\python38-32\lib\site-packages
\pandas\plotting_matplotlib\tools.py:298: MatplotlibDeprecationWarning:
The colNum attribute was deprecated in Matplotlib 3.2 and will be removed t
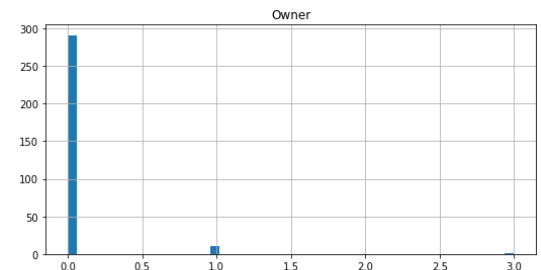wo minor releases later. Use ax.get_subplotspec().colspan.start instead.
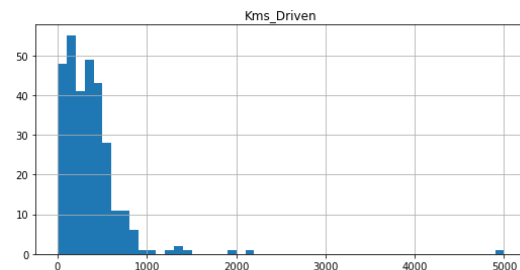 layout[ax.rowNum, ax.colNum] = ax.get_visible()
c:\users\santh\appdata\local\programs\python\python38-32\lib\site-packages
\pandas\plotting_matplotlib\tools.py:304: MatplotlibDeprecationWarning:
The rowNum attribute was deprecated in Matplotlib 3.2 and will be removed t
wo minor releases later. Use ax.get_subplotspec().rowspan.start instead.
  if not layout[ax.rowNum + 1, ax.colNum]:
c:\users\santh\appdata\local\programs\python\python38-32\lib\site-packages
\pandas\plotting_matplotlib\tools.py:304: MatplotlibDeprecationWarning:
The colNum attribute was deprecated in Matplotlib 3.2 and will be removed t
wo minor releases later. Use ax.get_subplotspec().colspan.start instead.
  if not layout[ax.rowNum + 1, ax.colNum]:

In [11]: ▶
```python
data.to_csv (r'C:\\Users\santh\OneDrive\Documents\Spring 2020 Assignments\M
```